

*Original Research Article***Exploring Northeast African Metric Craniofacial Variation at the Individual Level: A Comparative Study Using Principal Components Analysis**

S.O.Y. KEITA*

National Human Genome Center at Howard University, Anthropology Department, Smithsonian Institute, Washington, DC 20060

ABSTRACT A principal components analysis was carried out on male crania from the northeast quadrant of Africa and selected European and other African series. Individuals, not predefined groups, were the units of study, while nevertheless keeping group membership in evidence. The first principal component seems to largely capture “size” variation in crania from all of the regions. The same general morphometric trends were found to exist within the African and European crania, although there was some broad separation along a cline. Anatomically, the second principal component captures predominant trends denoting a broader to narrower nasal aperture combined with a similar shape change in the maxilla, an inverse relation between face–base lengths (“projection”) and base breadths, and a decrease in anterior base length relative to base breadth. The third principal component broadly describes trends within Africa and Europe: specifically, a change from a combination of a relatively narrower face and longer vault, to one of a wider face and shorter vault; it shows the northeast quadrant Africans along a cline with the other Africans. Stated in relative terms, the northeastern Africans tend to exhibit narrower bases in relationship to more projecting faces, and broader nasal areas than Europeans, although there is range of variation. Relative to the other African groups, they have narrower nasal areas and narrower faces in relationship to vault length. The crania from the northeast quadrant of Africa collectively demonstrate the greatest pattern of overlap with both Europeans and other Africans. Variation was found to be high in all series but greatest in the African material as a whole. Individuals from different geographical regions frequently plotted near each other, revealing aspects of variation at the level of individuals that is obscured by concentrating on the most distinctive facial traits once used to construct “types.” The high level of African interindividual variation in craniometric pattern is reminiscent of the great level of molecular diversity found in Africa. These results, coupled with those of Y chromosome studies, may help generate hypotheses concerning the length of time over which recent craniometric variation emerged in Africa. *Am. J. Hum. Biol.* 16:679–689, 2004. © 2004 Wiley-Liss, Inc.

The northeast quadrant of Africa has provided a range of data to researchers interested in human diversity and modern human origins. Fossils indicate the Pleistocene presence of people with anatomically modern or near-modern traits in the Horn and their antecedents (Lahr, 1996; White et al., 2003). Some molecular data are interpreted as suggesting that this region contains populations having a subset of wider African variation and that early modern humans from the area provided the founders of non-African populations (Tishkoff et al., 1996). The results of other molecular studies on the Y chromosome are interpreted as indicating that the southern part of this region, namely, the Horn/East Africa, was a major source of population in the Nile Valley and northwest Africa after

the Last Glacial Maximum, with some migration into the Near East and southern Europe (see Underhill et al., 2001; Bosch et al., 2001).

In the history of anthropology and African studies the Horn–Nile basin region figures prominently. Groups from the area were postulated to have a special relationship with Europe and/or the Near East. In one view the Horn was the hypothesized source of the

*Correspondence to: S. Keita, National Human Genome Center at Howard University, Howard University Hospital Cancer Center #615, 2041 Georgia Avenue NW, Washington, DC 20060.

Received 20 March 2004; Revision received 10 June 2004; Accepted 11 August 2004

Published online in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/ajhb.20076

ancestors of nearly all Europeans except some Germans, who together with some Africans formed a taxon called "Eurafrican" (Sergi, 1901). In another view, the Horn was the postulated place of origin of the ancestors of supra-Saharan Africans, southern Europeans, and southwest Asians, but not most northern or central Europeans, and called the "Brown Race" initially (Smith, 1911, 1923). However, the theory that has had the greatest influence, and still persists in one form or another in some disciplines, was put forth by Seligman (1930), who stated that the peoples of the Horn-Nile basin and supra-Saharan regions were *primarily* the descendants of southwest Asian immigrants *into* Africa (i.e., were settler colonists), who arrived at some time in the unspecified past. These immigrants were called "Hamites" (a term not original to Seligman), and constructed as "dark pastoral Europeans" who allegedly brought "Hamitic" languages, narrow noses and faces, linear body builds, lighter skin coloration, and any "significant" cultural innovation to Africa, with the exception, interestingly, of agriculture (see Sanders, 1969; MacGaffey, 1966).

Another idea constructed by Seligman was that of the "True Negro," a foil used as a baseline for discussions of admixture, and thus the creation of numerous taxa (actually phenotypes) in Africa with varying degrees of "Hamitic" and "Negro" ancestry (no "True White" was ever defined to discuss Europe or Asia). "True Negro" became equated with "real African"; this notion has not totally disappeared.

Hamites were divided into eastern and northern taxa, with the Egyptians, Nubians, and Somali (and certain other eastern Africans) being placed in the eastern group. Amazigh (Berbers) (and others, e.g., Fulbe) were classified into the northern wing. In essence, the peopling of northern and parts of eastern Africa in the post-glacial period was interpreted as being primarily the product of settler colonists who came from *outside* of Africa and penetrated down to East Africa at some time in the past—but apparently after the advent of horticulture in Seligman's view. The biological aspects of this theory were largely based on assumptions about the genesis of facial morphology and ideas about what constituted a "real African," but as Hiernaux (1975) states, there had long been peoples with narrow noses and faces, and less prognathism in East Africa, and migration from Eurasia

was unnecessary to explain their presence. Hiernaux calls this morphological trend "Elongated African" and postulates that it is the indigenous product of selection in a hot-dry climate. (The stereotyped "African" trend can likewise be called "Broad African.") Although Seligman's "Hamitic hypothesis" as a biocultural theory has been successfully critiqued by linguists, historians, and some biological anthropologists, some concepts and terms associated with this construct can still be found in the works of various disciplines, especially with regard to supra-Saharan Africans. It is beyond the scope of this presentation to fully explore this issue.

The recent molecular findings, history of ideas, and geography prompt further exploration of biological variation in the Horn-Nile Valley region in relation to that found in Europe and other parts of Africa. Given the role of physiognomy in previous writings a study of craniofacial variation was undertaken which captures aspects of anatomy beyond the visibly striking traits most noted by earlier anthropologists, and used by them to erect racial taxa.

This article presents the results of a principal components analysis (PCA) of craniometric data from the Nile Valley-Horn region, other African areas, and Europe. The primary goal was to assess the anatomical basis of patterns of craniofacial variation along an African-European continuum, with a special focus on northeastern African material, and secondarily to examine the mapping of the individuals in relationship to each other and groups of origin. There was interest in whether or not there was a sharp boundary separating any of these groups from each other, but especially the bulk of the northeastern African crania that all came from series that were once labeled "Hamitic." In addition to this, the directions of overlap were also of interest in an assessment of diversity and relative phenetic affinity.

This analysis is different from one using multiple discriminant functions in that the crania are not analyzed in predefined groups, with a technique designed to maximize the between- to within-group variation. Rather, the operational taxonomic units are the individual crania. Using this approach the diversity of individuals could be mapped while retaining knowledge of their geographical and series of origin. Interregional overlap, as exhibited by the positions of individuals, is to be examined in the spirit of

exploratory data analysis (EDA) (see Tukey, 1977). The emphasis on individuals and variation is meant to help circumvent the limitations of analyses whose starting point is predefined groups, and to explore beyond group thinking in order to get a different view of variation. This approach to human variation is consistent with the idea that the population may be the last "bastion" of the type concept (Brace, 1964).

MATERIALS AND METHODS

The series of primary interest are from the Horn, specifically, modern Somalia (n = 41, ca. 1940s); Kerma, an ancient Sudanese site (n = 48, ca. 1900 BC); Badari, northern Upper Egypt, (n = 25, ca. 4400–4000 BC); and Gizeh, a late dynastic northern Egyptian cemetery which overlaps the Graeco-Roman period (n = 58, ca. 600–300 BC). Only males were used due to their greater numbers and generally better condition in the target series (there were almost no females in the series from the Horn).

The data for the comparative material (and the Gizeh series) are from the Howells' (1973) database. The series are designated Dogon (Mali, n = 47); Zulu (South Africa, n = 55); Teita (Kenya, n = 33); Norse (Norway, n = 55); Berg (Hungary, n = 56); and Zalavar (Hungary, n = 53).

The Badarian, Kerma, Horn, Teita, and Gizeh series are housed in the Duckworth Laboratory at Cambridge University, Cambridge, England. (See Howells, 1973, for details about the other series; the Teita and Gizeh data are also from the Howells' database.)

PCA was undertaken using the SAS statistical package (SAS Institute, Cary, NC, 1990). The technique was applied to the mass of crania treated as a single group (but while keeping the place of origin in evidence by labeling). PCA transforms the original variables into new compound variables, or components, which geometrically approximate the same space in reduced dimensions. The components are orthogonal (unrelated) vectors successively derived from either the correlation or dispersion matrix calculated from the raw data (Blackith and Reyment, 1971). These vectors denote, in a summary fashion, the "qualities" which distinguish the individuals under investigation. Multiple components are derived that progressively describe less of the variation that differentiates the

individuals for the measurement battery employed. Plots are generated which show the spread of the individuals along the axes of variation described by the components.

The problems with interpretation are well known. The first component captures the greatest amount of interindividual variation, and in biometric work this has been usually interpreted as a general growth or "size" vector, when the analyses have been confined to one "population" (see Joliceur and Mosimann, 1969); this is based on the correlation coefficients (of variables to components) having only positive values. Subsequent components are said to indicate shape differences, due to the "disjoint variation of the various characters" (Joliceur and Mosimann, 1969:352), and are indicated by coefficients having positive and negative values. There are exceptions to the general claims about the meaning of the different components. The crania of holocene *Homo sapiens* fall into a fairly restricted size range. Also, it is important to keep in mind that shape and size concepts are to be understood in the context of a particular analysis, and are sometimes difficult to disentangle, since aspects of shape are related to changes in size (Bookstein, 1989), especially below the species level.

All of the embryological regions of cranium are represented. The variables are listed with their correlation coefficients as derived from the PCA.

RESULTS

PCA is performed on individuals, and does not generate mean (centroid) values for groups like canonical discriminant function analysis. This means that, in order to more easily examine distribution and overlap, the *density* of individuals from the various cranial series and regions should be considered in relationship to each other. This can be accomplished by using the axes of ordination as noted. The PCA plots can be viewed as mappings, with differentially dense regions. By assessing the density in relationship to cranial origin, some idea of distinctiveness of the series when individuals are assessed can be obtained.

Only the first three components are considered. These account for 43.3% of the variance (Table 1). The plots (Figs. 1, 2) appear daunting to examine, but close inspection is revealing and rewarding. A standard Cartesian orientation to the coordinates is useful in

TABLE 1. Eigenvalues and variance

	PC 1	PC 2	PC 3
Eigenvalue	10.4878	6.0570	3.3719
Difference	4.4308	2.6851	0.5879
Proportion	0.2280	0.1317	0.0733
Cumulative	0.2280	0.3597	0.4330

interpretation: PC 1 is the abscissa and PC 2 the ordinate in Figure 1, and PC 2 the abscissa and PC 3 the ordinate in Figure 2. Using the origin (0,0) and the quadrants that would be formed by perpendicular lines through them are useful guides in the evaluation of patterns. The analysis of distributions can be accomplished by systematically focusing on the extremes of the borders of the plotted individuals, and then looking at the densities of the individuals from various regions.

First, an examination will be undertaken of the variable correlations with the principal components in order to assess the morphometric qualities or trends suggested by them (see Table 2). Next, an evaluation of

regional and series distributions will be undertaken.

The first axis (PC 1) has only positive correlation coefficients. While conceptualizing the variance on PC 1 (Fig. 1) as only indicating size difference is conceptually problematic, it appears that in this study it has some explanatory power, and is hence at least somewhat justified, while being problematic. If only correlation coefficients with values greater than 0.5 are considered, then the "size" difference between the individuals is primarily a function of vault and anterior base lengths, facial breadths, and lengths (radii) from the transmeatal axis to various points in the face. A lesser but important contribution to overall "size" is made by breadths of the base, and vertical dimensions within the face, which have slightly lower coefficients. Inspection of the plot contrasting PC 1 and PC 2 (Fig. 1) reveals that the crania from the three regions evince some spread along PC 1, indicating a range of "sizes" within the groups; for example, note the distribution of Berg (A) crania along PC 1.

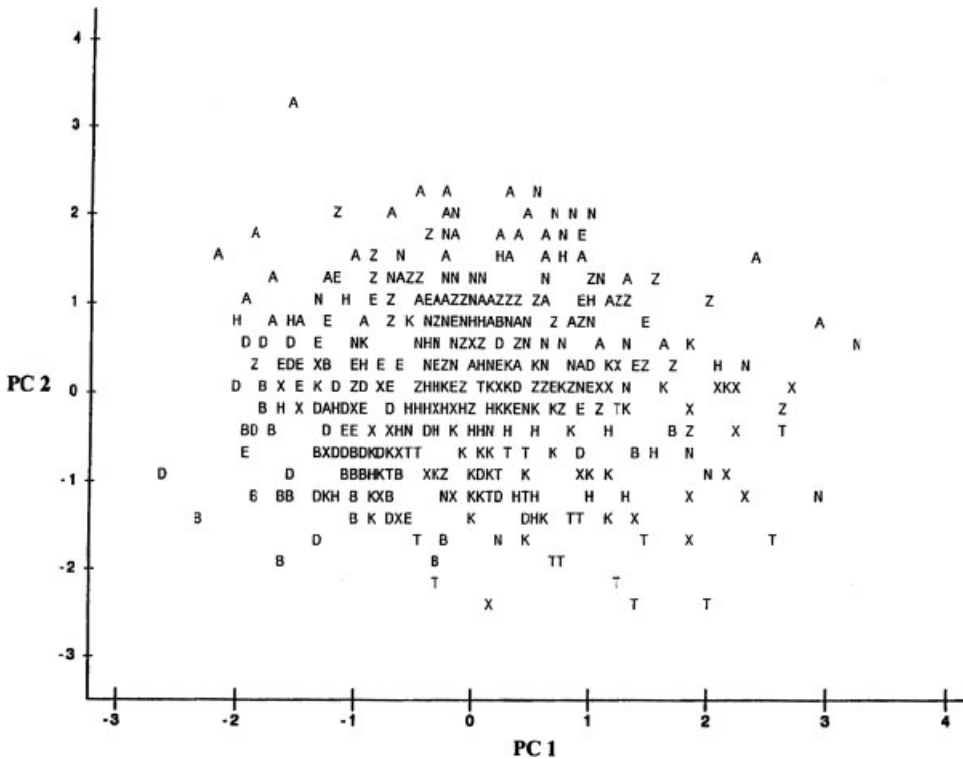


Fig. 1. Plot of the "size" (PC 1) versus "shape" (PC 2) components.

The coefficients of the second and third components have mixed signs, generally understood as more denoting of "shape" qualities. "Shape" contrasts, again accepting the conceptual limitations, indicated by PC 2 (Figs. 1, 2) include changes in the nasal aperture, maxilla, and base regions. There is also a contrast of shape described by base-face lengths in relationship to vault and base breadths. The major trends can be summarized as follows along PC 2 from its negative to positive values: 1) a broader to narrower nasal aperture combined with a similar shape change in the maxilla; 2) an inverse

relation between face-base lengths ("projection") and base breadths: i.e., greater to lesser facial projection with narrower to wider bases; and 3) a change (decrease) in anterior base length relative to base breadth.

The third component (Table 2), in the main, marks qualities of facial change including nasal shape in relation to vault length and height. Again, along the axis there is a change from a combination of a relatively narrower face and longer vault, to one of a wider face and shorter vault. There is also a general change from relatively narrower to broader nasal areas, but in relationship to

TABLE 2. Principal components: variables and associated correlation coefficients

	PC 1	PC 2	PC 3	
GOL	0.63632	0.11778	-0.44237	Glabella-occipital length
BNL	0.75496	-0.18625	-0.18845	Basion-nasion length
BBH	0.34933	0.17718	-0.37696	Basion-bregma height
XCB	0.18954	0.80767	0.01600	Maximum cranial breadth
XFB	0.28186	0.76392	0.12967	Maximum frontal breadth
STB	0.19304	0.74744	0.07507	Bistephanic breadth
ZYB	0.55949	0.47186	0.37444	Bizygomatic breadth
AUB	0.42558	0.66267	0.15806	Biauricular breadth
WCB	0.39965	0.49980	0.22613	Minimum cranial breadth
ASB	0.33890	0.63940	-0.04185	Biasterionic breadth
BPL	0.54937	-0.56633	0.19539	Basion-prosthion length
NPH	0.47134	0.37925	-0.08314	Nasion-prosthion height
NLH	0.43782	0.35451	-0.15763	Nasal height
OBH	0.10894	0.32488	0.15824	Orbit height
OBB	0.41617	0.33666	0.29384	Orbit breadth
NLB	0.24353	-0.32080	0.48905	Nasal breadth
MAB	0.46409	0.10250	0.41357	Palate breadth
ZMB	0.41675	-0.15123	0.37215	Bimaxillary breadth
SSS	0.22850	-0.02066	-0.15641	Zygomaxillary subtense
FMB	0.56779	0.15883	0.61452	Bifrontal breadth
NAS	0.39724	0.14146	-0.16555	Nasion-frontal subtense
EKB	0.56110	0.09880	0.62103	Biorbital breadth
DKS	0.18992	0.00328	-0.27760	Dacryon subtense
DKB	0.25814	-0.19892	0.36934	Interorbital breadth
NDS	0.24152	0.24979	-0.21816	Naso-dacryal subtense
IML	0.33228	-0.26850	0.08860	Inferior malar length
XML	0.49324	-0.03254	0.12568	Maximum malar length
MLS	0.11340	-0.23855	0.25090	Malar subtense
WMH	0.43302	0.20513	0.06519	Cheek height
FOL	0.30171	0.30053	0.06610	Foramen magnum length
FRC	0.31272	0.33349	-0.34865	Frontal chord
FRS	0.07335	0.17729	0.02310	Frontal subtense
PAC	0.26815	0.00210	-0.39723	Parietal chord
PAS	0.01620	0.13169	-0.31474	Parietal subtense
OCC	0.20643	0.15799	-0.32044	Occipital chord
OCS	0.25817	0.25084	-0.18222	Occipital subtense
VRR	0.24000	0.38670	-0.38813	Vertex radius
NAR	0.81963	-0.08152	-0.25739	Nasion radius
SSR	0.78325	-0.38712	-0.10471	Subspinale radius
PRR	0.68454	-0.51527	0.05461	Prosthion radius
DKR	0.80348	-0.17600	-0.19624	Dacryon radius
ZOR	0.79000	-0.38895	-0.12337	Zygoorbitale radius
FMR	0.69858	-0.16380	-0.19465	Frontomalare radius
EKR	0.74550	-0.29334	-0.12571	Ectoconchion radius
ZMR	0.71005	-0.43818	-0.05684	Zygomaxillare radius
AVR	0.74391	-0.43621	-0.00558	MI Alveolus radius

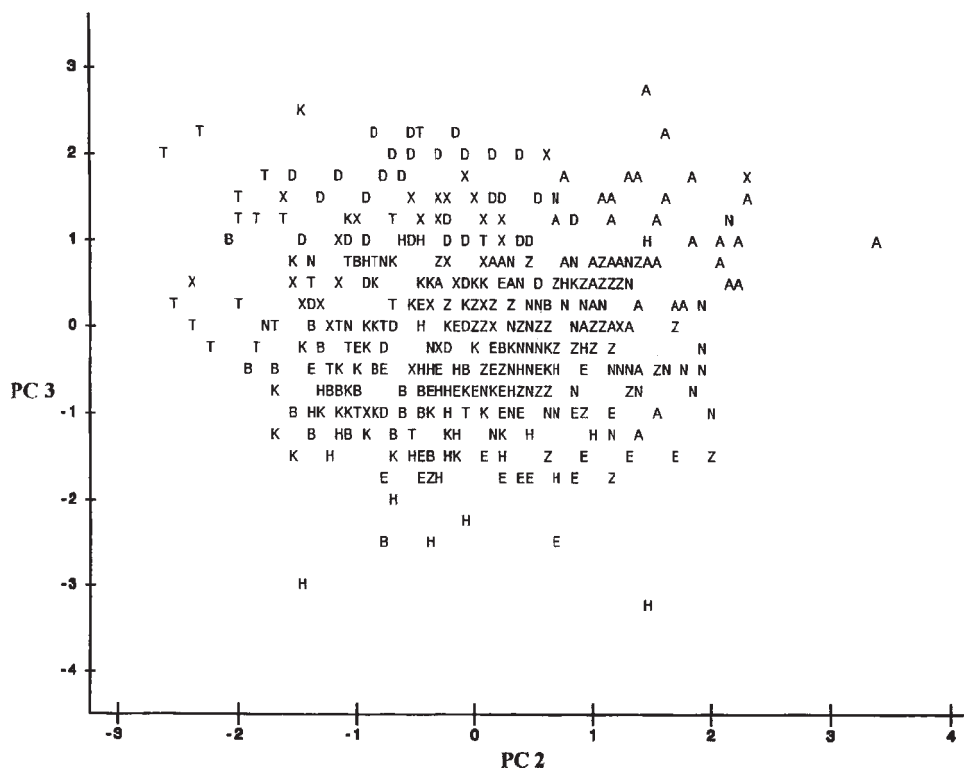


Fig. 2. Plot of the "shape" vectors PC 2 and PC 3.

the other variables; absolute size of nasal breadth may be little different and this is captured in relation to other patterns. Other trends of variation are more difficult to describe and include the shapes of the frontal, parietal, and occipital bones in relationship to anterior breadth of the neurocranium, with relative shortening of individual bone lengths in relation to widths along PC 3, as previously described.

In the "shape" plot of PC 2 by PC 3 (Fig. 2), some relative patterns emerge revealing a vague substructure related to the series origins, but again there is tremendous overlap. Inspection reveals few areas of clear-cut distinctions except at the extremes of the functions. There is a striking density with great heterogeneity near the center of the plot (0,0). Some continental trends are observable. Africans (Teita-T, Dogon-D, Zulu-X, Somali-H, Kerma-K, Badarian-B, Gizeh-E) and Europeans (Zalavar-Z, Berg-B, and Norse-N) (Table 3), broadly separate along PC 2, and this can be appreciated in both

Figures 1 and 2. However, there is extensive overlap in all directions, most notably by the African series. Within the continental "divisions" are the trends described by the anatomical correlates of PC 3.

The density of individuals from the European series, Berg (A), Norse (N), and Zalavar (Z), is greatest in the positive region of PC 2, further outlined by the positive and negative regions on PC 3: coordinates (0, -2), (2.5, -2.), (2.5,2), and (0,2). The lower half of this area contains more of the Norse crania and those from Zalavar. Individuals from the Berg series stand out in its upper half.

The African series as a whole concentrate into the negative region of PC 2 and also span the positive and negative area of PC 3, but as with the European crania, grade over into the other regions of the plot. The overall level of density is similar to that of the European series. The series traditionally thought of as "tropical": Dogon (D), Teita (T), and Zulu (X), map in greatest density into the region between the coordinates

TABLE 3. Series symbols

Badarian (predynastic Egypt)	B
Kerma (Kush, ancient Sudan)	K
Somali-Oromo ("Horn")	H
Egypt (Gizeh, late dynastic)	E
Dogon (Mali)	D
Norse (Norway)	N
Zalavar (Hungary)	Z
Berg (Hungary)	A
Teita (Kenya)	T
Zulu (South Africa)	X

(0,0), (0, 2), (-2.5,2) and (-2.5, 0). Note the relative distinctiveness of a subgroup of the Dogon (D), and how they plot within the same coordinates along PC 3 at the same points as Berg (A), but largely separated along PC 2. This position on PC 3 indicates a relative similarity for the composite metric quality represented by the morphometric trend on this axis of variation, but not "overall" similarity.

Individuals from the Nile Valley (Badarian-B, Kerma-K, Gizeh-E) and Horn (H) concentrate in the area of the plot bounded by coordinates (0,0), (0, -2), (-2, -2) and (0, -2) (Fig. 2, and incidentally 1 also). However, they also can be found in numerous areas; the Gizeh group (E) is notably diffuse; it overlaps notably with the region of the plot occupied by European material. PC 2 also describes trends within Africa and Europe, thus further highlighting the variation. The position of Nile Valley and Horn individuals considered collectively is seen to be at the "extreme" of the two morphometric trends, but overlapping the greatest with the other Africans and Europeans. Stated in relative morphometric terms, they tend to exhibit narrower bases in relationship to more projecting faces, and broader nasal areas than Europeans, although there is a range of variation. Relative to the other African groups they have narrower nasal areas, and narrower faces in relationship to vault length.

DISCUSSION

The plots are immediately striking in that sharp patterns of segregation of individuals by group origin do not emerge in the two-dimensional plots. The observable substructure, while challenging to interpret, is nevertheless interesting. Several points can be made. It is striking how much "size" varies by individual within the European and

African regions, assuming that PC 1 captures primarily this quality; Bergman's rule is not demonstrated in these data in any easily recognizable way, since individuals from all regions exhibit variation. Similar morphometric "shape" trends are found across regions and in a broad clinal distribution, with distinctiveness at the extremes of the axes of variation in individuals. Careful inspection reveals that the Horn and Nile Valley individuals (B, K, E, H) seem to show the most overlap with both other Africans and Europeans.

Although there are no striking boundaries in evidence, the fuzzy division along PC 2 into Africans and Europeans is noteworthy. Within these "divisions" are other metric pattern trends along PC 3. Both have nasal shape variation and the variability in the "ratio" of face breadth to vault length and height. However, collectively the African crania (X, B, T, E, D, H, K) would seem to show more variation than their European counterparts, in range and degree of variation. This is indicated by the degree and direction of scatter in the plots. Both the interindividual differences within the known series and the overlap of individuals from different regions are illustrative.

The concepts of similarity and variation at the individual level are made palpable by this analysis. The lack of uniformity of individuals from specific cranial series is starkly observable, although the distinctiveness of "extreme" subsets of individuals is apparent (e.g., some Berg (B), Teita (T), Dogon (D), and Norse (N)). It is clear that beyond the visually assessed "traits," or mean values/indices, and other traits traditionally used to define or describe types, are other aspects of variation that blur the boundaries between named groups. This degree of variation for one set of observations versus another is consistent with the findings of Relethford (2002), who in a study of data from various regions of the world found high within-group cranial variation as compared to skin color (one of those characters whose variants were used to make traditional classifications). In other words, the groups/entities "defined" by certain traits become more blurred when other measurements beyond those traditionally assessed are used, and many aspects of form measured, in an analysis of individuals, which of course is not a new discovery, but bears repeating. Relethford's (2002) global study is also interesting in this regard even at the group level,

because while some Nubians and Somali have skins just as darkly pigmented as some West and Central Africans, they generally have a different craniofacial pattern.

Groups of individuals "homogenized" by name (or selected cranio-facial traits) do not consistently emerge in this study. Hence, PC 3 also reveals subtle relative similarities (but not identities) between individuals from groups normally thought of as being "very different" in an absolute sense—for example, Berg (B) and Dogon (D). The relative resemblance of individuals from distant regions in aspects of multivariate craniofacial metric variation not traditionally recognized by small numbers of "traditional" measurements or "anatomical appreciation" cannot be overemphasized. The concept of individual variation for this trait battery is reinforced by the lack of emergence of clearly discrete two-dimensional clusters of individuals mirroring their series of origins, a point that bears repetition.

As noted, the interindividual African variability is perhaps greater when the degree and range of variation is considered compared with that of the Europeans. The Nile Valley and Horn groups show the greatest overlap with the other regions. This could mean that the northeast quadrant African patterns are more generalized (or that others are more specialized), and/or are more recent hybrids, or simply more variable, but none of these ideas can be definitively supported by this phenetic analysis. The overall results are generally consistent with findings of high African diversity, which in the main can now be considered to be primarily of indigenous African biohistorical origin, without denying some immigration with gene flow from various areas, especially southwest Asia. The paleo-anthropological and genetic data are more relevant in this regard, both at the levels of modern humans and dispersals after the late glacial maximum (see, e.g., White et al., 2003; Underhill et al., 2001). This interindividual variation in morphometrics when many variables are used indicates other avenues for enquiry in studies of phenetic variation. For example, to what factors is the patterning within the Sahel most related? The morphometric diversity of indigenous Africans (from Algeria to Zimbabwe) deserve wider recognition, and may help further explore human microevolution; this may be especially true for the variation in the northeast quadrant of Africa.

These results are interesting even only when the Africans are considered, and

especially in light of the history of ideas about biological classification of African peoples. However, the integration of these results with certain genetic data is even more intriguing. Recall that the Horn–Nile Valley crania show, as a group, the largest overlap with other regions. A review of the recent literature indicates that there are male lineage ties between African peoples who have been traditionally labeled as being "racially" different, with "racially" implying an ontologically deep divide. The PN2 transition, a Y chromosome marker, defines a lineage (within the YAP+ derived haplogroup E or III) that emerged in Africa probably before the last glacial maximum, but after the migration of modern humans from Africa (see Semino et al., 2004). This mutation forms a clade that has two daughter subclades (defined by the biallelic markers M35/215 (or 215/M35) and M2) that unites numerous phenotypically variant African populations from the supra-Saharan, Saharan, and sub-Saharan regions based on current data (Underhill, 2001).

The M2 lineage is mainly found primarily in "eastern," "sub-Saharan," and sub-equatorial African groups, those with the highest frequency of the "Broad" trend physiognomy, but found also in notable frequencies in Nubia and Upper Egypt, as indicated by the RFLP TaqI 49a, f variant IV (see Lucotte and Mercier, 2003; Al-Zahery et al. 2003 for equivalences of markers), which is affiliated with it. The distribution of these markers in other parts of Africa has usually been explained by the "Bantu migrations," but their presence in the Nile Valley in non-Bantu speakers cannot be explained in this way. Their existence is better explained by their being present in populations of the early Holocene Sahara, who in part went on to people the Nile Valley in the mid-Holocene, according to Hassan (1988); this occurred long before the "Bantu migrations," which also do not explain the high frequency of M2 in Senegal, since there are no Bantu speakers there either.

The 215/M35 lineage maps with the geographical arc primarily defined by the Horn and some areas south of it, down the Nile and over to Morocco (with a northern extension through the Levant into the Aegean), but its ancestral state is nearly confined to East Africa. This region also maps the core distribution of the Afro-Asiatic language family in Africa. The 215/M35 subclade has been further characterized with biallelic markers,

and found to have a group of daughter lineages of unique interest in Africa: M81 primarily found among Amazigh (Berber) speakers; and M78 found in East Africa and the Nile Valley among modern Egyptians (see the data in Underhill et al., 2001; Bosch et al., 2001; Cruciani et al., 2004; Luis et al., 2004; Semino et al., 2004; and in Lucotte et al., 2003, for TaqI p49a, f RFLP haplotypes V and XI which in Africa also signify M35/215). These data, considered together, make it possible to see these groups as being coextensive with each other, and therefore allow the extension and revision of Hiernaux's evolutionary model to include a range of ancestral supra-Saharan peoples or the major component in their male lineages. The idea of linking these populations from east Africa to Morocco, and postulating an African origin for them, is not new (see Angel and Kelley, 1986), but now receives lineage genetic support.

The issue of gene flow/colonization into the region and when it occurred deserves some mention given the history of ideas. In some analyses this could confuse the issue of core origins if all data are not considered. The YAP insertion that defines haplogroup DE has been said to be of "Asian" origin, which would mean that there had been a major migration from Asia (see Templeton, 2002); however, the genetic evidence (Underhill and Roseman, 2001) and archaeological data best support an African origin. The evidence for a very ancient migration to Africa from the Near East, and evincing gene flow, but not replacement, is the presence of the M89/M9/M173 in the R haplogroup, found primarily in the west Central African region and to a lesser degree in Egypt (Kittles, pers. commun.). In any case, migrations before the demonstrable emergence of known language families and food production belong to a level of biological history that long antedates the interaction of current ethnic groups, and care should be taken not to express admixture in terms of these.

Migration and gene flow at a later time is an issue. There is the ongoing concern, which either parallels or is a remake of the Hamitic theory, of migration from Asia to Africa in the form of demic diffusion; this model is sometimes offered to explain the presence of agriculture and/or the Afro-Asiatic language family in Africa. This will only be briefly discussed. While it is probable that there was some migration from the Near East during the Neolithic, along with the spread of ideas, neither the genetic nor the linguistic evidence

support a demic diffusion model. It is noteworthy that neither the Berber nor ancient Egyptian language families are branches of Semitic, Indo-European, or Sumerian. Given this, it is more interesting that the terms for the key Near Eastern domesticates (wheat, barley, sheep, goat) are not even Semitic loans in ancient Egyptian (Takacs, pers. commun.), even though horticulture with ovacaprines in Africa is first attested in northern Egypt—some 2,000 years after it was established in the Near East! Finally, the cautious interpretation of the evidence on ancestral Afro-Asiatic indicates that it was spoken by non-food producers, and emerged in Africa in the Horn or southeastern Sahara (Bender, 1975; Blench, 1993; Diakonoff, 1981, 1998; Ehret, 1984, 1995; Fleming, 1974, pers. commun.; Greenberg, 1966, 1973; Newman, pers. commun.; Nichols, 1997). The Y genetic profiles of the Horn–Nile Valley region are different from those of the core Semitic-speaking populations in the Near East, who are characterized by a high frequency of M89 variants in the J and R haplogroups. (But as previously noted, the M35 lineage was taken into the Near East before the Neolithic perhaps by pre-proto-Semitic speakers).

Lower Egypt most certainly has been the recipient of immigrants in ancient as well as more recent times (Lucotte and Mercier, 2003). The delta region of Egypt has been impacted by European (Graeco-Roman) and Near Eastern peoples, the latter apparently primarily during the Islamic and not Neolithic period (Nebel et al., 2002). It is important to say that the indigenous northern Egyptians, while adjacent to the Libyco-Berber region, cannot simply be called "Berbers." The Y chromosome data suggest that the original Egyptian Nile Valley population cannot be treated analytically as "Berber," thereby in effect negating the distinctiveness and identity of the core indigenous Nile Valley populations (see, e.g., Harich et al., 2002; Luis et al., 2004; Herrera et al., 2004, for a description of "Egyptians" as merely being an "Arab"–"Berber" admixture/composite, without a discussion of the indigenous Nile Valley population).

Some Horn populations assimilated southwest Asians, and even adopted their languages, which likely began as lingua francas. Certain Ethiopian groups evince substantial frequencies of "Near Eastern" genes (Y chromosome J group lineages), likely due to the assimilation of migrants after the first

millennium BC (Munro-Hay, 1991), with some founder effect, but this is not substantially true for Oromo and Somali peoples (see, e.g., Comas et al., 1999; Sistonen et al., 1987). Linguistic evidence does not suggest that Semitic speakers brought agriculture to Ethiopia. (The Ethiopian genetic profile may have valid alternative explanations incorporating *bi-directional* migrations and settlements of great antiquity, depending on how old linguistics would predict the ancestral Ethio-Semitic language to be, in order to account for the present linguistic variation.) Ancient gene flow from such migrations would have been reworked by the new environment and demographic factors, and thus become a part of African biological history. It is important to say that there is no evidence to suggest that in the Holocene population *replacement* occurred in any of these regions as a whole based on the Y chromosome data. Populations should be viewed processually as dynamic entities over time and not "static" entities. The presence of M35/215 lineages and the Benin sickle cell variant in southern Europe illustrates this well.

The consideration of the results of this analysis, with the observation that the African peoples of the PN2 clade are tremendously diverse in skin color, body build, facial conformation, hair form, etc., forces a reexamination of the ideas about the relationship of craniofacial similarity to descent group in Africa. All of the African populations represented in this study, or related peoples, are either in the PN2 or YAP+ "family" at some substantial level. From a general evolutionary perspective, the PN2 clade trenchantly demonstrates the point that individuals/populations having a tremendous range of physical traits, the focus of traditional approaches, can have fairly recent common ancestry in terms of lineage, i.e., can derive substantively from a common population, and therefore, in this sense, be more related than they are to others to whom they may be more phenetically similar in craniometric pattern (or classical polymorphisms). This Y-lineage reveals (micro) cladistic relationships not predicted by the traditional racial approach, which cut across phena, and true breeding populations as well. In short, phenetic affinity is shown to not be concordant with Y-cladistic affinity for the level of African biohistory represented by these markers.

The results of this work and that of Relethford (2002), in the context of the distribution of the PN2 clade, can be used to gen-

erate hypotheses about the time in which a range of hard and soft part anatomical variation can evolve in a Y chromosome-defined descent group. The phylogeographic approach advocated by Avise (2000) thus shows its potential importance in appreciating some axes and pathways of human variation in more restricted locales. A modified version of the climatic hypotheses of Hiernaux (1975), coupled with considerations of drift/founder effect in a metapopulation model involving colonization and desiccation cycles in the fragmented habitat of the Holocene Sahara, and other regions, likely help explain the dispersion of the subclades of PN2 and differential survival of various sublineages in various populations. For the history of ideas, the diversity in the craniometric patterns in northeastern Africa, and the distribution of PN2 daughter lineages, render as falsified aspects of Seligman's "Hamitic hypothesis" and subsequent hypotheses consciously or unconsciously based on it. For future work in the morphometrics of the northeastern quadrant of Africa, the patterns noted here and the genetic data may suggest new ways to consider the complexity of variation in regionally defined human groups, and to consciously understand populations as processes, not static entities.

ACKNOWLEDGMENTS

I thank Dr. R. Foley, Dr. Marta Lahr, and Mrs. M. Bellatti for facilitated access to the collections of the Duckworth Laboratory. Professor W.W. Howells kindly shared his data. Dr. A.J. Boyce provided some useful perspectives on data exploration, and the notion of "mapping diversity." Discussions with Drs. P. Underhill, O. Semino, C. Mulligan, R. Kittles, C. Ehret, H. Fleming, P. Newman, R. Blench, and G. Takacs were invaluable. I also thank three anonymous referees who helped improve this article. Mrs. M. Evans typed the initial manuscript.

LITERATURE CITED

- Al-Zahery N, Semino O, Benuzzi G, Magri C, Passarino G, Torroni A, Santachiara-Benerecetti AS. 2003. Y-chromosome and mtDNA polymorphisms in Iraq, a crossroad of the early human dispersal and of post-Neolithic migrations. *Mol Phylogenet Evol* 28:458-472.
- Angel L, Kelley J. 1986. Description and comparison of the skeleton. In: Wendorf F, Schild R, editors. *The Wadi Kubanniya skeleton*. Dallas: Southern Methodist University Press. p 53-70.
- Avise JC. 2000. *Phylogeography*. Cambridge, MA: Harvard.

- Bender ML. 1975. Omotic: A New Afro-Asiatic Language Family. Carbondale: University Museum, Southern Illinois University.
- Blench R. 1993. Recent developments in African language classification and their implications for prehistory. In: Shaw T, Sinclair P, Andah B, and Okpoko A, editors. *The Archaeology of Africa*. London: Routledge. p 126–138.
- Blackith RE, Reyment RA. 1971. *Multivariate morphometrics*. London: Academic Press.
- Brace CL. 1964. On the race concept. *Curr Anthropol* 5:313–314.
- Bookstein FL. 1989. “Size and shape”: a comment on semantics. *Syst Zool* 38:173–189.
- Bosch E, Calafell F, Comas D, Oefner P, Underhill PA, Bertranpetit J. 2001. High-resolution analysis of human Y-chromosome variation shows a sharp discontinuity and limited gene flow between northwestern Africa and the Iberian peninsula. *Am J Hum Genet* 68:1019–1029.
- Comas D, Reynolds R, Sajantila A. 1999. Analysis of mtDNA HVRII in several human populations using an immobilized SSO probe hybridization assay. *Eur J Hum Genet* 7:459–68.
- Cruciani F, La Fratta R, Santolamazza P, Sellitto D, Pascone R, Moral P, Watson E, Guida V, Colomb EB, Zaharova B, Lavinha J, Vona G, Aman R, Cali F, Akar N, Richards M, Torroni A, Novelletto A, Scozzari R. 2004. Phylogeographic analysis of haplogroup E3b (E-215) Y chromosomes reveals multiple migratory events within and out of Africa. *Am J Hum Genet* 74:1014–1022.
- Diakonoff (Diakonov) IM. 1981. Earliest Semites in Asia. *Altorientalische Forschungen* 8:23–74.
- Diakonoff (Diakonov) IM. 1998. The earliest Semitic society. *Linguistic data. J Semitic Studies* 4:209–19.
- Ehret C. 1984. Historical/linguistic evidence for early African food production. In: Clark, JD, Brandt S, editors. *From hunters to farmers*. Berkeley: University of California Press. p 26–36.
- Ehret C. 1995. *Reconstructing Proto-Afroasiatic (Proto-Afrasian): vowels, tone, consonants and vocabulary*. University of California Publications in Linguistics, 126. Berkeley: University of California Press.
- Fleming H. 1974. Omotic as an Afroasiatic family. *Studies in African Linguistics*, suppl 5:81–94.
- Greenberg JH. 1966. *The Languages of Africa*. Bloomington: Indiana University Press.
- Greenberg JH. 1973. African languages. In: Skinner EP, editor. *Peoples and Cultures of Africa*. Garden City, New York: Doubleday Natural History Press. p 34–58.
- Harich N, Esteban E, Chafik A, Lopez-Alomar A, Vona G, Moral P. 2002. Classical polymorphisms from Moyen Atlas (Morocco): genetics, geography, and historical evidence in the Mediterranean peoples. *Ann Hum Biol* 29:473–487.
- Herrera RJ, Adrien LR, Ruiz LM, Sanabria NY, Duncan G. 2004. D1S80 single-locus discrimination among African populations. *Hum Biol* 76:87–108.
- Hiernaux J. 1975. *The people of Africa*. New York: Charles Scribner’s Sons.
- Howells WW. 1973. *Cranial variation in Man*. Cambridge, MA: Peabody Museum.
- Joliceur P, Mosimann JE. 1960. Size and shape variation in the painted turtle. A principal component analysis. *Growth* 24:339–354.
- Keita SO. 1990. Studies of ancient crania of northern Africa. *Am J Phys Anthropol* 83:245–254.
- Lahr MM. 1996. *The evolution of modern human diversity: a study of cranial variation*. Cambridge, UK: Cambridge University Press.
- Lucotte G, Mercier G. 2003. Y-chromosome haplotypes in Egypt. *Am J Phys Anthropol* 121:63–66.
- Luis JR, Rowold DJ, Regueiro M, Caeiro C, Cinnoglu C, Roseman C, Underhill PA, Cavalli-Sforza LL, Herrera RJ. 2004. The Levant versus the Horn of Africa: evidence for bi-directional corridors of human migrations. *Am J Hum Genet* 74:532–544.
- MacGaffey W. 1966. Concepts of race in the historiography of northeast Africa. *J Afr Hist* 7:1–17.
- Munro-Hay S. 1991. *Aksum: An African civilization of late antiquity*. Edinburgh: Edinburgh University.
- Nebel A, Landau-Tasserot E, Filon D, Oppenheim A, Faerman M. 2002. Genetic evidence for the expansion of Arabian tribes into the southern Levant and North Africa. *Am J Hum Genet* 70:1594–1596.
- Nichols J. 1997. Modeling ancient population structures and movement in linguistics. *Annual Review of Anthropology* 26:359–384.
- Relethford JH. 2002. Apportionment of global human genetic diversity based on craniometrics and skin color. *Am J Phys Anthropol* 118:393–398.
- Sanders E. 1969. The Hamitic hypothesis: its origin and functions in time perspective. *J Afr Hist* 10:521–532.
- Seligman C. 1930. *The races of Africa*. Oxford: Oxford University Press (see also 1966 edition).
- Semino O, Magri C, Benuzzi G, Lin A, Al-Zahery N, et al. 2004. Origin, diffusion, and differentiation of Y-chromosome haplogroups E and J: inferences on the neolithization of Europe and later migratory events in the Mediterranean area. *Am J Hum Genet* 74:1023–1034.
- Sergi G. 1901. *The Mediterranean race*. London: Walter Scott.
- Sistonen P, Koistinen J, Aden Abdulle O. 1987. Distribution of blood groups in the East African Somali population. *Hum Hered* 37:300–313.
- Smith GE. 1911. *The ancient Egyptians and the origin of civilization*. London: University College (2nd ed., 1923).
- Templeton A. 2002. Out of Africa again and again. *Nature* 416:45–51.
- Tishkoff SA, Dietzsch E, Speed W, Pakstis AJ, Kidd JR, Cheung K, and Bonne-Tamir B. 1996. Global patterns of linkage disequilibrium at the CD4 locus and modern human origins. *Science* 271:1380–1387.
- Underhill P, Roseman CC. 2001. The case for an African rather than an Asian origin of the human Y-chromosome YAP insertion. In: Jin L, Seielstad M, Xiao C, editors. *Genetic, linguistic and archaeological perspectives on human diversity in southeast Asia*. London: World Scientific. p 43–56.
- Underhill PA, Passarino G, Lin AA, Shen P, Lahr MM, Foley RA, Oefner P, Cavalli-Sforza LL. 2001. The phylogeography of Y chromosome binary haplotypes and the origins of modern human populations. *Ann Hum Genet* 65:43–62.
- White TD, Asfaw, DeGusta D, Gilbert H, Richards GD, Suwa G, Howell FC. 2003. Pleistocene *Homo sapiens* from Middle Awash, Ethiopia. *Nature* 423:742–747.